PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-083257
(43)Date of publication of application : 31.03.1998

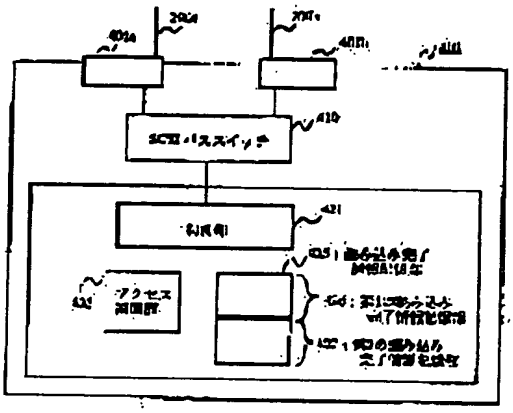
(51)Int.Cl. G06F 3/06
G06F 9/445
G06F 11/16

(21)Application number : 08-238013
(22)Date of filing : 09.09.1996

(71)Applicant : MITSUBISHI ELECTRIC CORP
(72)Inventor : ITO KAZUHIKO
MIZUNO MASAHIRO
YAMAMOTO HITOSHI

(54) DATA STORAGE SYSTEM AND DATA STORAGE MANAGING METHOD

(57)Abstract:
PROBLEM TO BE SOLVED: To obtain a data storage system which can start two host systems cooperatively without a conflict in a data storage system shared by two host system.
SOLUTION: A decision part 421 discriminates which host system gains access. An access control part 423 exclusively control the access to a disk by the host system decided by the decision part 421 so that no conflict is caused. Further, a read completion information storage part 425 makes the two host systems know the completion of the system starting mutually.



LEGAL STATUS

[Date of request for examination] 30.10.1996
[Date of sending the examiner's decision of rejection]
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]
[Date of final disposal for application]
[Patent number] 2830857
[Date of registration] 25.09.1998
[Number of appeal against examiner's decision of rejection]
[Date of requesting appeal against examiner's

(19)日本国特許庁（ＪＰ）　(12)公開特許公報（Ａ）　(11)特許出願公開番号

特開平10－83257

(43)公開日　平成10年(1998)３月31日

(51)Int.Cl.⁶		識別記号	庁内整理番号	ＦＩ		技術表示箇所
Ｇ０６Ｆ	3/06	３０１		Ｇ０６Ｆ	3/06	３０１Ｃ
	9/445				11/16	３１０Ｚ
	11/16	３１０			9/06	４２０Ｋ

審査請求　有　請求項の数14　ＯＬ　（全 13 頁）

(21)出願番号　特願平8－238013

(22)出願日　平成８年(1996)９月９日

(71)出願人　000006013
三菱電機株式会社
東京都千代田区丸の内二丁目２番３号

(72)発明者　伊藤　一彦
東京都千代田区丸の内二丁目２番３号　三菱電機株式会社内

(72)発明者　水野　正博
東京都千代田区丸の内二丁目２番３号　三菱電機株式会社内

(72)発明者　山本　整
東京都千代田区丸の内二丁目２番３号　三菱電機株式会社内

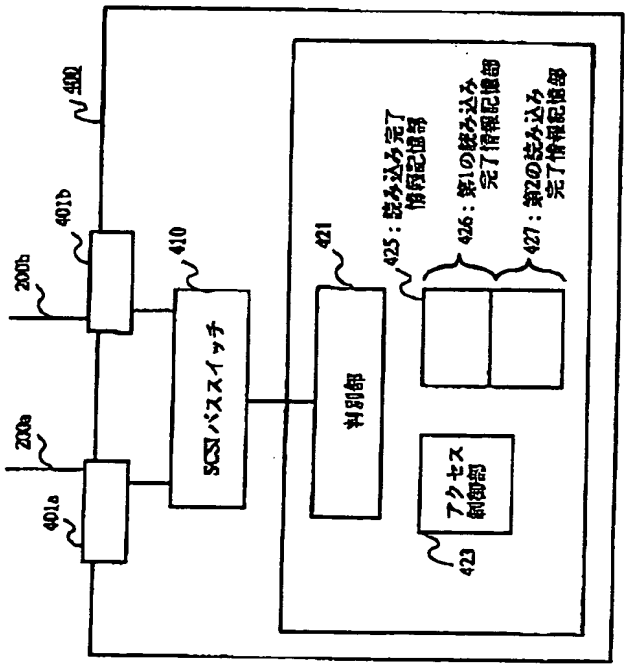
(74)代理人　弁理士　宮田　金雄　（外３名）

(54)【発明の名称】　データストレージシステム及びデータストレージ管理方法

(57)【要約】

【課題】　２つのホストシステムで共有されるデータストレージシステムにおいて、２つのホストシステムに競合を発生させないように、協調してシステム起動を行えるデータストレージシステムを得る。

【解決手段】　判別部４２１がどちらのホストシステムからのアクセスかを判別する。アクセス制御部４２３は、判別部４２１により判別されたホストシステムのディスクへのアクセスを排他的に制御し、競合を発生させない。また、読み込み完了情報記憶部４２５により、２つのホストシステムは、お互いにシステム起動の完了を知る。



【特許請求の範囲】

【請求項１】　所定のオペレーティングシステムで動作する第１と第２のホストシステムに接続され、上記第１と第２のホストシステムからアクセスされるデータストレージシステムにおいて、

以下の要素を有するデータストレージシステム

（ａ）所定のインタフェースを持つ第１と第２のパス、

（ｂ）上記第１のパスを介して上記第１のホストシステムに接続され、上記第２のパスを介して上記第２のホストシステムに接続されるとともに上記所定のオペレーティングシステムを記憶するストレージ、（ｃ）上記第１のパスを介して上記第１のホストシステムに接続され、上記第２のパスを介して上記第２のホストシステムに接続され、上記第１と第２のホストシステムの起動時に上記第１と第２のホストシステムによる上記ストレージに記憶されたオペレーティングシステムの読み込みを排他的に行わせるストレージマネージャ。

【請求項２】　上記ストレージマネージャは、上記第１と第２のホストシステムの起動時に、上記ストレージより先に上記第１と第２のホストシステムからアクセスされるように構成され、上記ストレージマネージャは、一方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込みを優先させて実行させ、他方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込みを待たせるとともに、

上記一方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込み完了後に、上記他方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込みを実行することにより上記第１と第２のホストシステムのオペレーティングシステムの読み込みを排他的に行わせることを特徴とする請求項１記載のデータストレージシステム。

【請求項３】　上記ストレージマネージャは、上記第１と第２のパスに接続され上記第１と第２のパスを切換ることにより上記第１と第２のパスのいずれかを選択的に接続すると共に、上記第１と第２のパスのどちらが接続されているかを示すパス接続情報を出力するパス切換機構と、

上記パス切換機構が出力するパス接続情報を参照し、上記第１と第２のホストシステムのどちらからアクセスされたかを判別する判別部と、

上記判別部の判別に基づき、上記第１と第２のホストシステムの上記ストレージへのアクセスを排他制御し、上記ストレージに記憶されたオペレーティングシステムの読み込みを排他的に行うアクセス制御部を備えたことを特徴とする請求項２記載のデータストレージシステム。

【請求項４】　上記アクセス制御部は、上記第１と第２のパスのいづれか一方のパスをフリーズすることを特徴とする請求項３記載のデータストレージシステム。

【請求項５】　上記ストレージマネージャは、少なくとも上記第１と第２の一方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込みの完了を示す読み込み完了情報を記憶する読み込み完了情報記憶部を備え、上記アクセス制御部は、上記読み込み完了情報記憶部を参照して、上記一方のホストシステムのオペレーティングシステムの読み込み完了後に上記ストレージへのアクセスの排他制御を解除することを特徴とする請求項３に記載のデータストレージシステム。

【請求項６】　上記読み込み完了情報記憶部は、第１と第２の読み込み完了情報記憶部を備え、上記第１と第２の読み込み完了情報記憶部は、上記第１と第２のホストシステムのオペレーティングシステムの読み込みの完了を示す読み込み完了情報をそれぞれ個別に記憶するとともに、上記アクセス制御部は、第１と第２のホストシステムのいずれか先にアクセスがあったホストシステムのオペレーティングシステムの読み込みの完了まで、他方の後にアクセスのあったホストシステムのオペレーティングシステムの読み込みを待たせることを特徴とする請求項５に記載のデータストレージシステム。

【請求項７】　上記ストレージマネージャは、上記アクセス制御部により上記第１と第２のパスのうち一方のパスをフリーズすることにより、上記第１と第２のホストシステムのうちフリーズしていない他方のパスに接続されたホストシステムに上記ストレージに記憶されたオペレーティングシステムを読み込む優先権を与え、

上記他方のパスに接続されたホストシステムは、オペレーティングシステムの読み込みの完了時に上記第１と第２の読み込み完了情報記憶部の一方に読み込み完了情報を記憶し、

上記アクセス制御部は、上記一方の読み込み完了情報記憶部を参照して、上記一方のパスのフリーズを解除することを特徴とする請求項６に記載のデータストレージシステム。

【請求項８】　上記ストレージは、ディスク装置であり、上記所定のインタフェースは、スモール・コンピュータ・システム・インタフェース（ＳＣＳＩ）であり、上記アクセス制御部は、スモール・コンピュータ・システム・インタフェース（ＳＣＳＩ）のディスク起動コマンドに対する応答を遅延することによりパスをフリーズすることを特徴とする請求項４記載のデータストレージシステム。

【請求項９】　所定のオペレーティングシステムの読み込みにより起動する第１と第２のホストシステムと上記オペレーティングシステムを記憶するデータストレージシステムとが接続され、上記データストレージシステムがストレージとストレージマネージャを有しているデータストレージシステムのデータストレージ管理方法において、

以下の工程を有するデータストレージ管理方法

（ａ）上記第１と第２のホストシステムのいずれか一方のホストシステムが上記ストレージからオペレーティングシステムを読み込む工程、（ｂ）上記ストレージマネージャが他方のホストシステムの上記ストレージへのアクセスをロックするアクセスロック工程、（ｃ）上記一方のホストシステムによるオペレーティングシステムの読み込みの完了を検出し、アクセスロック工程による他方のホストシステムのアクセスのロックを解除するロック解除工程、（ｄ）ロックを解除された他方のホストシステムが上記ストレージからオペレーティングシステムの読み込みを行う工程。

【請求項１０】　上記アクセスロック工程は、（ａ）他方のホストシステムの上記ストレージへのアクセス要求を検出するアクセス要求検出工程、（ｂ）上記アクセス要求検出工程により検出されたアクセス要求に対する応答を遅延する応答遅延工程、を有することを特徴とする請求項９記載のデータストレージ管理方法。

【請求項１１】　上記データストレージ管理方法は、更に、上記一方のホストシステムのオペレーティングシステムの読み込み完了を示す読み込み完了情報を記憶する読み込み完了情報記憶工程を有し、上記ロック解除工程は、上記読み込み完了情報記憶工程により記憶された読み込み完了情報を参照して、他方のホストシステムのアクセスのロックを解除することを特徴とする請求項９記載のデータストレージ管理方法。

【請求項１２】　第１と第２のホストシステムに接続され、上記第１と第２のホストシステムからアクセスされるデータストレージシステムにおいて、

以下の要素を有するデータストレージシステム

（ａ）所定のインタフェースを持つ第１と第２のバス、

（ｂ）上記第１のバスを介して上記第１のホストシステムに接続され、上記第２のバスを介して上記第２のホストシステムに接続されるとともにデータを記憶するストレージ、（ｃ）上記第１と第２のホストシステムの上記ストレージに対するアクセスの可否を示すアクセス情報を予め設定するアクセス情報設定部、（ｄ）上記アクセス情報設定部に設定されたアクセス情報を参照し、上記第１のホストシステムの上記ストレージに対するアクセスを制御する第１のアクセス制御部、（ｅ）上記アクセス情報設定部に設定されたアクセス情報を参照し、上記第２のホストシステムの上記ストレージに対するアクセスを制御する第２のアクセス制御部。

【請求項１３】　上記データストレージシステムは、更に上記アクセス情報を変更するアクセス情報変更部を備え、

上記第１と第２のアクセス制御部の少なくともいずれかは、一定の時間間隔で上記アクセス情報を参照することにより、上記第１と第２のホストシステムのうち少なくともいずれかのホストシステムの上記ストレージに対するアクセスを動的に制御することを特徴とする請求項１２に記載のデータストレージシステム。

【請求項１４】　上記データストレージシステムは、上記第１と第２のホストシステムの少なくともいずれか一方のホストシステムの障害を検出する障害検出部を備え、上記アクセス情報変更部は、上記障害検出部が障害を検出した一方のホストの上記ストレージへのアクセスを禁止すると共に、他方のホストの上記ストレージへのアクセスを許可するよう上記アクセス情報を変更することを特徴とする請求項１３記載のデータストレージシステム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】この発明は、２つのホストシステムで共有されるデータストレージシステムに関するものである。また、２つのホストシステムで共有されるデータストレージ管理方法に関するものである。

【０００２】

【従来の技術】

従来例１．図１３は、１つのディスクサブシステムを２つのホストシステムで共有するシステムの構成図である。図１３において、１００ａは現用系（主系）のホストシステム、１００ｂは予備系（従系）のホストシステムである。１１０ａはホストシステム１００ａのディスク制御装置、１１０ｂはホストシステム１００ｂのディスク制御装置である。ディスク３００は、ホストシステム１００ａ，１００ｂで共有される複数のディスクである。２００ａは、ディスク３００とホストシステム１００ａを接続するバス、２００ｂは、ディスク３００とホストシステム１００ｂを接続するバスである。このような構成を採るシステムの一例として、デュプレクスシステム（Ｄｕｐｌｅｘ　Ｓｙｓｔｅｍ）がある。デュプレクスシステムは、スタンドバイシステムとも呼ばれ、オンライン処理を行う現用系と予備系のプロセッサから構成される。予備系は、通常バッチ処理などの優先度の低い業務に使用される。現用系が障害になると予備系の業務を中止し、障害装置を予備に切り換えて処理を再開する。予備系を常時サービス開始可能な状態に待機させておく方式をホットスタンドバイシステムと呼ぶ（電子情報通信ハンドブック、電子情報通信学会編、１９８８年３月３０日発行による）。また、予備系をサービス開始可能な状態に待機させておかない方式もある。この方式をコールドスタンドバイシステム（Ｃｏｌｄ　Ｓｔａｎｄ　ｂｙ　ｓｙｓｔｅｍ）と呼ぶ。コールドスタンドバイシステムの場合には、現用系が障害になった時点で、ＯＳ（オペレーティングシステム）の読み込み等の起動操作を行い、サービス開始可能な状態とする。即ち、コールドスタンドバイシステムの場合には、現用系ホストシステムの起動がまず行われ、現用系が障害になった時点で、予備系ホストシステムの起動が行われる。このため、それぞれのホストシステムの起動時には、ホストシ

ステムとディスクサブシステムが1対1となるので通常のシステムと同様に起動動作が行われる。一方、ホットスタンドバイシステムの場合には、共有しているディスクサブシステムに対して主系・従系ともにアクセスを行うため、特に起動時に以下の問題が発生する。その問題とは、ディスクサブシステムに対して従系がアクセスしている間に主系のアクセスが受け付けられず、主系が動作を中断してハングアップ（HungUp）してしまう可能性があることである。なぜならば、システム起動中は、ホストシステムがBIOS（Basic Input Output System）レベルで動作している為、OSの下で動作する本来のディスク制御用のドライバがロードされているときのようにリトライ動作等が十分に行われない為である。上記の理由から、デュプレクスシステム構成をとることが難しいという問題点があった。また、デュプレクスシステム構成でも従系をコールドスタンドバイにするなどの機能的な縮退が必要であるという問題点があった。また、デュプレクスシステムで従系をホットスタンドバイとする為には、主系、従系の間で協調してシステム起動を行うような方式をとる必要があった。

【0003】従来例2．次に、上記従来例1と同一の構成のシステムにおける他の問題点について述べる。主系（または従系）に障害が生じた場合、両系による運転から従系（または主系）のみの運転に切り替えるシステムにおいて以下の問題が存在する。第1の問題は、障害が生じた主系（または従系）の運転をシャットダウンする場合に、障害が生じた主系（または従系）がデータストレージにもアクセスするため、同じデータストレージをアクセスしている従系（または主系）の運転に干渉する可能性があるという点である。第2の問題は、従系（または主系）のみの運転中（縮退運転中）、メンテナンスの為に主系（または従系）を立ち上げた場合に、従系（または主系）の運転に干渉する可能性があるという点である。このように、ディスクサブシステムを共有する一方の系の動作が他の系に影響を及ぼし、場合によっては、システムダウンに陥る可能性があるという問題点があった。

【0004】

【発明が解決しようとする課題】この発明は、上記のような問題点を解決するためになされたものであり、2つのホストシステムで共有されるデータストレージシステムにおいて、システム起動時に主系・従系の間で競合を発生させない様に協調してシステム起動を行えるデータストレージシステム及びその管理方法を得ることを目的としている。また、2つのホストシステムで共有されるデータストレージシステムにおいて、2つのホストシステムの切換時及び一方のホストシステムの保守時のデータストレージへの干渉を防ぐ機構を備えたデータストレージシステムを得ることを目的としている。

【0005】

【課題を解決するための手段】この発明に係るデータストレージシステムは、所定のオペレーティングシステムで動作する第1と第2のホストシステムに接続され、上記第1と第2のホストシステムからアクセスされるデータストレージシステムにおいて、以下の要素を有することを特徴とする。

（a）所定のインタフェースを持つ第1と第2のバス、

（b）上記第1のバスを介して上記第1のホストシステムに接続され、上記第2のバスを介して上記第2のホストシステムに接続されるとともに上記所定のオペレーティングシステムを記憶するストレージ、（c）上記第1のバスを介して上記第1のホストシステムに接続され、上記第2のバスを介して上記第2のホストシステムに接続され、上記第1と第2のホストシステムの起動時に上記第1と第2のホストシステムによる上記ストレージに記憶されたオペレーティングシステムの読み込みを排他的に行わせるストレージマネージャ。

【0006】上記ストレージマネージャは、上記第1と第2のホストシステムの起動時に、上記ストレージより先に上記第1と第2のホストシステムからアクセスされるように構成され、上記ストレージマネージャは、一方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込みを優先させて実行させ、他方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込みを待たせるとともに、上記一方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込み完了後に、上記他方のホストシステムの上記ストレージに記憶されたオペレーティングシステムの読み込みを実行することにより上記第1と第2のホストシステムのオペレーティングシステムの読み込みを排他的に行わせることを特徴とする。

【0007】上記ストレージマネージャは、上記第1と第2のバスに接続され上記第1と第2のバスを切換ることにより上記第1と第2のバスのいずれかを選択的に接続すると共に、上記第1と第2のバスのどちらが接続されているかを示すバス接続情報を出力するバス切換機構と、上記バス切換機構が出力するバス接続情報を参照し、上記第1と第2のホストシステムのどちらからアクセスされたかを判別する判別部と、上記判別部の判別に基づき、上記第1と第2のホストシステムの上記ストレージへのアクセスを排他制御し、上記ストレージに記憶されたオペレーティングシステムの読み込みを排他的に行うアクセス制御部を備えたことを特徴とする。

【0008】上記アクセス制御部は、上記第1と第2のバスのいづれか一方のバスをフリーズすることを特徴とする。

【0009】上記ストレージマネージャは、少なくとも上記第1と第2の一方のホストシステムの上記ストレー

ジに記憶されたオペレーティングシステムの読み込みの完了を示す読み込み完了情報を記憶する読み込み完了情報記憶部を備え、上記アクセス制御部は、上記読み込み完了情報記憶部を参照して、上記一方のホストシステムのオペレーティングシステムの読み込み完了後に上記ストレージへのアクセスの排他制御を解除することを特徴とする。

【００１０】上記読み込み完了情報記憶部は、第１と第２の読み込み完了情報記憶部を備え、上記第１と第２の読み込み完了情報記憶部は、上記第１と第２のホストシステムのオペレーティングシステムの読み込みの完了を示す読み込み完了情報をそれぞれ個別に記憶するとともに、上記アクセス制御部は、第１と第２のホストシステムのいずれか先にアクセスがあったホストシステムのオペレーティングシステムの読み込みの完了まで、他方の後にアクセスのあったホストシステムのオペレーティングシステムの読み込みを待たせることを特徴とする。

【００１１】上記ストレージマネージャは、上記アクセス制御部により上記第１と第２のバスのうち一方のバスをフリーズすることにより、上記第１と第２のホストシステムのうちフリーズしていない他方のバスに接続されたホストシステムに上記ストレージに記憶されたオペレーティングシステムを読み込む優先権を与え、上記他方のバスに接続されたホストシステムは、オペレーティングシステムの読み込みの完了時に上記第１と第２の読み込み完了情報記憶部の一方に読み込み完了情報を記憶し、上記アクセス制御部は、上記一方の読み込み完了情報記憶部を参照して、上記一方のバスのフリーズを解除することを特徴とする。

【００１２】上記ストレージは、ディスク装置であり、上記所定のインタフェースは、スモール・コンピュータ・システム・インタフェース（ＳＣＳＩ）であり、上記アクセス制御部は、スモール・コンピュータ・システム・インタフェース（ＳＣＳＩ）のディスク起動コマンドに対する応答を遅延することによりバスをフリーズすることを特徴とする。

【００１３】この発明に係るデータストレージ管理方法は、所定のオペレーティングシステムの読み込みにより起動する第１と第２のホストシステムと上記オペレーティングシステムを記憶するデータストレージシステムとが接続され、上記データストレージシステムがストレージとストレージマネージャを有しているデータストレージシステムのデータストレージ管理方法において、以下の工程を有することを特徴とする。

（ａ）上記第１と第２のホストシステムのいずれか一方のホストシステムが上記ストレージからオペレーティングシステムを読み込む工程、（ｂ）上記ストレージマネージャが他方のホストシステムの上記ストレージへのアクセスをロックするアクセスロック工程、（ｃ）上記一方のホストシステムによるオペレーティングシステムの読み込みの完了を検出し、アクセスロック工程による他方のホストシステムのアクセスのロックを解除するロック解除工程、（ｄ）ロックを解除された他方のホストシステムが上記ストレージからオペレーティングシステムの読み込みを行う工程。

【００１４】上記アクセスロック工程は、（ａ）他方のホストシステムの上記ストレージへのアクセス要求を検出するアクセス要求検出工程、（ｂ）上記アクセス要求検出工程により検出されたアクセス要求に対する応答を遅延する応答遅延工程、を有することを特徴とする。

【００１５】上記データストレージ管理方法は、更に、上記一方のホストシステムのオペレーティングシステムの読み込み完了を示す読み込み完了情報を記憶する読み込み完了情報記憶工程を有し、上記ロック解除工程は、上記読み込み完了情報記憶工程により記憶された読み込み完了情報を参照して、他方のホストシステムのアクセスのロックを解除することを特徴とする。

【００１６】この発明に係るデータストレージシステムは、第１と第２のホストシステムに接続され、上記第１と第２のホストシステムからアクセスされるデータストレージシステムにおいて、以下の要素を有することを特徴とする。

（ａ）所定のインタフェースを持つ第１と第２のバス、（ｂ）上記第１のバスを介して上記第１のホストシステムに接続され、上記第２のバスを介して上記第２のホストシステムに接続されるとともにデータを記憶するストレージ、（ｃ）上記第１と第２のホストシステムの上記ストレージに対するアクセスの可否を示すアクセス情報を予め設定するアクセス情報設定部、（ｄ）上記アクセス情報設定部に設定されたアクセス情報を参照し、上記第１のホストシステムの上記ストレージに対するアクセスを制御する第１のアクセス制御部、（ｅ）上記アクセス情報設定部に設定されたアクセス情報を参照し、上記第２のホストシステムの上記ストレージに対するアクセスを制御する第２のアクセス制御部。

【００１７】上記データストレージシステムは、更に上記アクセス情報を変更するアクセス情報変更部を備え、上記第１と第２のアクセス制御部の少なくともいずれかは、一定の時間間隔で上記アクセス情報を参照することにより、上記第１と第２のホストシステムのうち少なくともいずれかのホストシステムの上記ストレージに対するアクセスを動的に制御することを特徴とする。

【００１８】上記データストレージシステムは、上記第１と第２のホストシステムの少なくともいずれか一方のホストシステムの障害を検出する障害検出部を備え、上記アクセス情報変更部は、上記障害検出部が障害を検出した一方のホストの上記ストレージへのアクセスを禁止すると共に、他方のホストの上記ストレージへのアクセスを許可するよう上記アクセス情報を変更することを特徴とする。

【００１９】
【発明の実施の形態】
実施の形態１．以降の実施の形態では、ストレージがディスクであり、ストレージマネージャがディスクマネージャである場合を例にとって説明する。また、ディスクサブシステムがホストシステムにＳＣＳＩ（Ｓｍａｌｌ Ｃｏｍｐｕｔｅｒ Ｓｙｓｔｅｍ Ｉｎｔｅｒｆａｃｅ）バスで接続される場合について説明する。図１は、この発明のデータストレージシステムの構成図である。図１に示すように、この実施の形態の構成は、図１３に示した従来の構成にディスクマネージャ４００を加えたものである。また、ディスク３００は、主系・従系両方のホストシステム１００ａ，１００ｂにそれぞれロードされるオペレーティングシステム３１０を記憶している。また、この実施の形態の構成は、主系・従系の２つのホストシステムにそれぞれ接続される２本のＳＣＳＩバス２００ａ，２００ｂを備えている。他の構成については、図１３に示した従来の構成と同様である。また、ディスクサブシステムを共有する２つのホストシステムは、デュプレクスシステムを構成することを想定しているが、２つのホストシステムが同一動作を行うデュアルシステムでも構わない。ディスクマネージャ４００は、共用ディスクサブシステム上に接続されるディスクサブシステム管理機構である。ディスクマネージャ４００は、ディスク３００と同様に主系・従系の両方のホストシステムからアクセス可能なデバイスである。また、ディスクマネージャ４００は、ホストシステムからは、ディスクと同等に見える。以下、主系のホストシステムを単に主系ともいう。また、従系のホストシステムを単に従系ともいう。
【００２０】図２は、ディスクマネージャ４００の構成図である。ディスクマネージャ４００は、主系・従系の２つのホストシステムにそれぞれ接続される２本のＳＣＳＩバス２００ａ，２００ｂに接続されている。ＳＣＳＩバスは、それぞれ主系用ポート４０１ａ，従系用ポート４０１ｂを経て、ＳＣＳＩバススイッチ４１０に接続される。このように、ディスクマネージャ４００は、必要に応じて２つのＳＣＳＩバスを切り替えるデュアルポートディスク（Ｄｕａｌ Ｐｏｒｔ Ｄｉｓｋ）の構成を有する。ＳＣＳＩバススイッチ４１０は、ＳＣＳＩバスコントロール４２０に接続されている。また、ディスクマネージャ４００は、ＭＰＵ（Ｍｉｃｒｏ Ｐｒｏｃｅｓｓｉｎｇ Ｕｎｉｔ）４４０と、不揮発性のメモリ４３０を有している。このように、ディスクマネージャ４００は、ＳＣＳＩバス上からは、ディスクとして認識され、ディスクと同等の動作を実行することが可能な構成となっている。ＳＣＳＩバス上からアクセスされるデータは、ディスクマネージャ４００上のメモリ４３０のＳＣＳＩブロックエリア４３２に記録される。メモリ４３０は、ディスクマネージャプログラムエリア４３８を有している。ディスクマネージャプログラムエリア４３８は、ＭＰＵ４４０で実行されディスクマネージャ４００を制御するディスクマネージャプログラムを記憶する。
【００２１】図３は、この発明のデータストレージシステムの機能ブロック図である。図３に示す判別部４２１は、バス切換機構であるＳＣＳＩバススイッチ４１０からのポート番号を受け取り、受け取ったポート番号からどちらのバスからアクセスされたかを判別する。これにより、どちらのホストシステムからアクセスされたかが判別される。アクセス制御部４２３は、判別部４２１により判別されたホストシステムのディスクへのアクセスを制御する。読み込み完了情報記憶部４２５は、第１の読み込み完了情報記憶部４２６と、第２の読み込み完了情報記憶部４２７からなる。アクセス制御部４２３は、後述するように、この読み込み完了情報記憶部４２５を参照して、各ホストシステムのディスクへのアクセスを制御する。
【００２２】次に、図４から図６を用いて、動作について具体的に説明する。この実施の形態では、システムに電源を投入する際、まず、主系のホストシステムから先に電源を投入し、次に、従系のシステムに電源を投入するものとする。また、主系、従系のシステム構成は、同一の構成であり、電源投入後のシステム起動にかかる時間もそれぞれ、ほぼ同一であることを想定している。また、ディスクマネージャ４００は、ディスク３００よりも先にホストシステム１００ａ，１００ｂからアクセスされる構成となっているものとする。図４は、主系のホストシステムの、電源投入からシステム起動完了までの流れ図である。図５は、主系と従系のホストシステムの電源投入からシステム起動完了までの、ディスクマネージャの動作の流れ図である。図６は、従系のホストシステムの、電源投入からシステム起動完了までの流れ図である。図４に示すように、まず主系がシステムのハードウェアへの電源投入（Ｓ１１０）でシステム起動を開始する。次に、Ｓ１２０において、主系のホストシステムは、主系システムのハードウェア（Ｈ／Ｗ）をチェックする。チェック終了後、入出力制御装置上のＢＩＯＳ（Ｂａｓｉｃ Ｉｎｐｕｔ Ｏｕｔｐｕｔ Ｓｙｓｔｅｍ）によりディスクのリードライト（Ｒｅａｄ／Ｗｒｉｔｅ）動作を開始する（Ｓ１３０）。まず、主系のホストシステムは、ディスクサブシステムに対して、リセット（Ｒｅｓｅｔ）を発行する（Ｓ１３２）。次に、主系のホストシステムは、接続されているディスクを確認するコマンド（Ｔａｒｇｅｔ Ｓｅｌｅｃｔ，Ｉｎｑｕｉｒｙ ＣＭＤ，Ｔｅｓｔ Ｕｎｉｔ Ｒｅａｄｙ）を発行する（Ｓ１３４）。そして、主系のホストシステムは、ディスクのバス上をディスクマネージャから順番にサーチして、接続されているディスクそれぞれに対してディスクの識別子であるＩＤの所定の順（昇順または降

順）にディスク起動コマンド（ＳｔａｒｔＣＭＤ）を発行していく。ディスクマネージャ４００は、ホストシステムからはディスクとして認識される。また、ディスクマネージャ４００は、ディスクよりも先に各ホストシステムからアクセスされる構成となっている。この為、主系のホストシステム１００ａは、まず最初にディスクマネージャ４００にディスク起動コマンドを発行する（Ｓ１３６）。

【００２３】一方、従系も、主系に続いて、システムのハードウェアへの電源投入でシステム起動を開始する（図６、Ｓ３１０）。次に、Ｓ３２０において、従系のホストシステムは、従系システムのハードウェア（Ｈ／Ｗ）をチェックする。チェック終了後、入出力制御装置上のＢＩＯＳ（Ｂａｓｉｃ　Ｉｎｐｕｔ　Ｏｕｔｐｕｔ　Ｓｙｓｔｅｍ）によりディスクのリードライト（Ｒｅａｄ／Ｗｒｉｔｅ）動作を開始する（Ｓ３３０）。まず、従系のホストシステムは、ディスクサブシステムに対して、リセット（Ｒｅｓｅｔ）を発行する（Ｓ３３２）。この時、すでに主系がディスクからＯＳの読み込みを実行中であれば、ディスクサブシステムに対して主系のホストシステムに接続されているパス２００ａが有効となる。主系のパス２００ａが有効である間、従系のパス２００ｂからのアクセスは無視される。従って、ディスクサブシステムは、従系から発行されたリセットを無視する。このため、従系のリセット発行により、主系のＯＳ読み込みが妨げられることはない。次に、従系のホストシステムは、接続されているディスクを確認するコマンドを発行する（Ｓ３３４）。そして、従系のホストシステムは、ディスクのパス上をディスクマネージャから順番にサーチして、接続されているディスクそれぞれに対してディスクの識別子であるＩＤの所定の順（昇順または降順）にディスク起動コマンドを発行していく。ディスクマネージャ４００は、ホストシステムからはディスクとして認識される為、従系のホストシステムは、主系と同様に、まず最初にディスクマネージャ４００にディスク起動コマンドを発行する（Ｓ３３６）。

【００２４】ディスクマネージャ４００は、主系・従系両方のホストシステムからディスク起動コマンドを受け取る。受け取ったディスク起動コマンドには、コマンドが主系用ポート，従系用ポートの内、どちらのポートから入ってきたかを示すポート番号がついている。ディスクマネージャ４００の判別部４２１は、このポート番号を参照して、主系・従系のうちどちらからディスク起動コマンドが発行されたかを判別してアクセス制御部４２３へ出力する。アクセス制御部４２３は、ディスク起動コマンドが主系から発行された場合には、そのディスク起動コマンドに対して応答する。図５のＳ２１０に示すように、ディスクマネージャ４００は、主系からのディスク起動コマンドを受け取ると、主系に対して、起動完了を報告する。また、Ｓ２２０に示すようにディスク起動コマンドが従系から発行された場合には、ディスク起動コマンドを受け取った時点で従系のディスクのパス２００ｂをフリーズ（ＦＲＥＥＺＥ）する。パス２００ｂがフリーズされると、従系はディスク起動コマンドの完了応答待ちになる（Ｓ３３７）。一般にディスク起動コマンドに対する応答のタイムアウトは設定されない為、従系は応答があるまで無制限に待つことになる。主系は、ディスクマネージャ４００のアクセス制御部からディスク起動コマンドに対する起動完了を受け取ると、次に接続されているディスク３００に対してディスク起動コマンドを発行する。その後、主系は、ディスクからＯＳ（オペレーティングシステム）の読み込み（ロード）を行う（Ｓ１３８）。ＯＳの読み込みが完了すると、主系のホストシステムのシステム起動が完了する。主系は、システムの起動完了後、ディスクマネージャ４００に対して起動完了を報告する（Ｓ１４０）。起動完了の報告は、主系のホストシステムからディスクマネージャの管理するＳＣＳＩブロックエリアに完了コード（起動完了フラグ）を記録することで実行される。

【００２５】図７は、ディスクマネージャのＳＣＳＩブロックエリアを示す図である。主系・従系ともにディスクマネージャ４００上のＳＣＳＩブロックを共有する為、以下の方法で起動の完了報告を受け取ることができる。ＳＣＳＩブロック０は、主系からは、リードライト（読み書き）可能なＳＣＳＩブロックであり、従系からはリード（読み込み）専用のＳＣＳＩブロックである。また、ＳＣＳＩブロック１は従系からは、リードライト（読み書き）可能なＳＣＳＩブロックであり、主系からはリード（読み込み）専用のＳＣＳＩブロックである。主系・従系それぞれのホストシステムは、これらのＳＣＳＩブロックのうち書き込みする権利のあるＳＣＳＩブロックに起動完了フラグ４３３ａ，４３４ａの書き込みを行う。他の系は、起動完了フラグの書き込みをチェックすることでシステム起動完了を検出できる。主系には、ディスクマネージャのＳＣＳＩブロックエリアの主系のホストシステム用に定められた所定の場所であるＳＣＳＩブロック０に対して、完了コード（例えば、数字の１）の書き込みを行う書き込み手段（プログラム）が用意されている。そして、その書き込み手段は、システムの起動完了後に自動的に起動されるよう予め設定されているものとする。その設定に従い、書き込み手段は、完了コードをディスクマネージャのＳＣＳＩブロックエリアの主系のホストシステム用に定められた所定の場所であるＳＣＳＩブロック０に記録する。主系のホストシステムは、その後、適当な時間間隔でＳＣＳＩブロック１をポーリングしながら、従系のシステム起動完了を待つ（Ｓ１５０）。

【００２６】アクセス制御部４２３は、一定の時間間隔で、ディスクマネージャのＳＣＳＩブロックエリアの主系のホストシステム用に定められた所定の場所（ＳＣＳ

Ｉブロック０）をチェックしている。アクセス制御部は、そのチェックにより、主系のシステム起動が完了したことを確認し（Ｓ２３０）、従系のパスのフリーズを解除し、ディスク起動コマンドに対する応答を返す（Ｓ２４０）。

【００２７】従系は、ディスクマネージャ４００のアクセス制御部４２３からディスク起動コマンドに対する起動完了を受け取ると、接続されているディスク３００に対してディスク起動コマンドを発行する。その後、従系は、ディスクからＯＳ（オペレーティングシステム）の読み込み（ロード）を行う（Ｓ３３８）。ＯＳの読み込みが完了すると、従系のホストシステムのシステム起動が完了する。従系は、システムの起動完了後、ディスクマネージャ４００に対して起動完了を報告する（Ｓ３４０）。従系の起動完了の報告も主系と同様に、従系のホストシステムからディスクマネージャの管理するＳＣＳＩブロックエリアに完了コードを記録することで実行される。主系と同様に、従系には、ディスクマネージャのＳＣＳＩブロックエリアの従系のホストシステム用に定められた所定の場所（ＳＣＳＩブロック１）に対して、完了コード（例えば、数字の１）の書き込みを行う書き込み手段が用意されている。そして、その書き込み手段は、主系の書き込み手段と同様に、システムの起動完了後に自動的に起動されるよう予め設定されているものとする。その設定に従い、書き込み手段は、完了コードをディスクマネージャのＳＣＳＩブロックエリアの従系のホストシステム用に定められた所定の場所（ＳＣＳＩブロック１）に記録する。ＳＣＳＩブロック１をポーリングしている主系のホストシステムは、完了コードの書き込みによりシステム起動完了を知る。その後、主系と従系は、ディスクサブシステムを共有して、デュプレクスシステムの運用を開始する。

【００２８】なお、ここでは、主系から発行されたディスク起動コマンドが先にディスクマネージャに受け取られる場合について、図５に示している。これは、前述したように、主系と従系のシステムが同一構成であり、主系のホストシステムの電源投入が先に行われるため、主系のディスク起動コマンドが先に発行されることを想定しているためである。だが、どちらのホストシステムからコマンドが発行されたかは、前述したように判別部により判別されるため、仮に従系のディスク起動コマンドが先に発行されても構わない。即ち、図５のＳ２１０とＳ２２０は順不同で構わない。従系のディスク起動コマンドが先に発行された場合には、従系のパスが主系のディスク起動コマンドの発行に先立ちフリーズされる。その後、主系からディスク起動コマンドが発行されると、主系に対して、起動完了が報告される。その後のＯＳの読み込み動作は、前述した場合と同様である。

【００２９】次に、ディスクマネージャの動作について説明する。図８は、システム起動時のディスクマネージャの動作を示す流れ図である。ディスクマネージャ４００は、通常は主系・従系の動作に対する応答をＳＣＳＩバスを通じて行う。システム起動時には図８に示すシーケンスで動作を実行する。まず、Ｓ５００において、ディスクマネージャ４００の初期化が行われる。次に、Ｓ５１０において主系・従系両方のシステム起動が完了しているかをチェックする。完了していれば、Ｓ５２０において、通常動作に移行する。完了していなければ、Ｓ５３０において、主系のシステム起動が完了しているかどうかをチェックする。主系のシステム起動が完了していなければ、Ｓ５４０において、主系のコマンドに対する応答を優先して行う。このとき、従系からのディスク起動コマンドが発行されたら、アクセス制御部により、従系のＳＣＳＩバスをフリーズする。主系のシステム起動が完了していれば、Ｓ５５０において、起動コマンドをフリーズしているかどうか判定する。フリーズしていれば、Ｓ５６０において、起動コマンドに対する応答を実行する。フリーズしていなければ、従系は、コマンドの応答待ちの状態ではないので、Ｓ５６０の処理はスキップする。次に、Ｓ５７０において、従系のコマンドに対する応答を優先して行う。

【００３０】このように、アクセス制御部は、システム起動時の主系と従系のディスクアクセスを制御し、ＯＳの読み込みを排他的に行わせる。それにより、ＯＳの読み込みで主系と従系の競合が発生しなくなる。競合が発生しないのでＢＩＯＳによるＯＳの読み込み時にエラーの発生を防ぐことができる。

【００３１】実施の形態２．この実施の形態では、２重系のディスクに対する競合を制御するシステム管理機構を備えたデータストレージシステムについて説明する。図９は、この実施の形態のデータストレージシステムの構成図である。ディスクマネージャ４００は、ＳＣＳＩ等のディスクインタフェースを有するディスクサブシステム管理機構（ハードウェア）である。ディスクマネージャ４００は、主系と従系のホストシステムの各ディスクへのアクセスの許可（○）／禁止（×）をアクセス情報として登録するディスク管理テーブル１４００を有している。アクセス制御部１２０ａ，１２０ｂはディスク管理テーブル１４００を参照してホストシステム１００ａ，１００ｂのディスクへのアクセスを制御する。アクセス情報変更部１２２ａ，１２２ｂは、ディスク管理テーブル１４００に登録されているアクセス情報を必要に応じて変更する。また、障害検出部１２４ａ，１２４ｂは、ホストシステム１００ａ，１００ｂに発生した障害を検出する。図１０に、ディスク管理テーブルの一例を示す。図１０に示すディスク管理テーブルでは、主系のホストシステムがディスク＃０にアクセスを許可され（１４１０）、従系のホストシステムがディスク＃１にアクセスを許可されている（１４２０）。ここでは、アクセス情報として、アクセス許可とアクセス禁止を

“○”と“×”で示しているが、他の値でも構わない。主系のディスク制御装置がディスク＃０、従系のディスク制御装置がディスク＃１を占有して主系と従系が動作する。このアクセスの可否を示すアクセス情報は、システム構築時にディスクマネージャのディスク管理テーブルへセットされる。ハードウェアの構成からはどちらのディスク制御装置からもディスク＃０、＃１共アクセス可能であるが、アクセスを許可されたディスク以外はアクセスしないように、アクセス制御部が制御を行なう。

【００３２】アクセス制御部１２０ａ，１２０ｂの動作について説明する。ディスク管理テーブル１４００には、システムの起動前に予めアクセス情報が設定されているものとする。ディスクマネージャ４００は、ＳＣＳＩ等のディスクインタフェースを有する。アクセス制御部１２０ａ，１２０ｂはこのディスクインタフェースを介してディスク管理テーブル１４００にアクセスする事ができる。また、ディスクマネージャ４００は、デュアルポート機構を持ち、２つの系からアクセスが可能である。アクセス制御部１２０ａ，１２０ｂは、ディスク管理テーブルを参照して、その系に対するディスクのアクセス許可または禁止を判別し、アクセスが禁止されているディスクへのアクセスを行なわない様に制御する。このように、ディスク管理テーブルを用いることで、物理的には、アクセス可能な複数のディスクを、ホストシステム毎に占有して使用することができる。

【００３３】ディスク管理テーブルのアクセス情報は、アクセス情報変更部１２２ａ，１２２ｂにより必要に応じて随時変更される。一方、アクセス制御部１２０ａ，１２０ｂは、一定間隔でディスクマネージャ上のディスク管理テーブルを参照して、アクセス情報をチェックする。それにより、一旦システム起動した後でも、ディスクへのアクセスの制御（許可または禁止）をダイナミックに切り替えることができる。ディスク管理テーブルのアクセス情報の変更について具体的に説明する。システムに障害が生じた場合、ディスクアクセスを行なわずに安全（他系に影響をおよぼさないよう）にシャットダウン動作を行なう為に、以下のようにディスク管理テーブルのアクセス情報を書き換える。ここでいうシステムに生じた障害とは、例えば、ローカルエリアネットワーク（ＬＡＮ：Ｌｏｃａｌ　ＡｒｅａＮｅｔｗｏｒｋ）のネットワーク用制御カード（図示せず）に発生したトラブルなど、ホストシステムのＣＰＵは正常に動作し、また、ホストシステムのアクセス制御部や、アクセス情報変更部及び障害検出部は動作可能であるような障害であるものとする。ここでは、障害が主系のホストシステム１００ａに発生した場合について説明する。障害検出部１２４ａは、ホストシステム１００ａに発生した障害を検出する。障害検出部１２４ａは障害が発生したことをアクセス情報変更部に知らせる。アクセス情報変更部は、主系のホストシステム１００ａからのアクセスが許可になっているディスク＃０のアクセス情報をアクセス禁止に書き換えて、主系のホストシステム１００ａのディスク＃０へのアクセスを禁止とすると共に、従系のホストシステム１００ｂからのディスク＃０へのアクセスを許可するように、ディスク管理テーブル１４００のディスク＃０のアクセス情報を書き換える。図１１に、アクセス情報変更後のディスク管理テーブルを示す。図１１の、１４３０に示すように主系からのディスク＃０に対してはアクセス禁止となり、１４４０に示すように従系からディスク＃０に対してはアクセス許可となっている。従系のアクセス制御部１２０ｂがディスク管理テーブルをポーリングして、ディスク＃０が“アクセス禁止”から“アクセス許可”に変更になった事を知る。主系は、システムシャットダウン時、及び、診断用に再起動した場合も含めて、ディスク管理テーブルを参照する事によって、ディスクへのアクセスが禁止されている事を知り、ディスクへアクセスすることはない。従って、主系が従系の動作に干渉する事はない。

【００３４】次に、主系のホストシステム１００ａがダウンし、主系のアクセス制御部１２０ａ、アクセス情報変更部１２２ａ、障害検出部１２４ａが動作しない場合について説明する。この場合の、システムダウンの検出は、従系の障害検出部１２４ｂが行う。図１２は、この一方の系の障害検出部が他方の系の障害を検出する場合のシステム構成図である。図１２において、１１００は、システムが生きているというパルスであるハートビートを主系と従系で送受信するハートビート信号線である。ハートビートは、システムが生きている間、一定の時間間隔で発生するパルスである。このハートビートを監視することにより、一定の時間間隔でハートビートが確認できれば、システムが正常稼働していることが判り、ハートビートが確認できなければ、なんらかの障害が発生したことを検出できる。障害検出部１２４ｂは、ハートビート信号線１１００からハートビートを入力し監視することにより、主系のホストシステムのダウンを検出する。システムダウンを検出すると、障害検出部１２４ｂは、アクセス情報変更部１２２ｂに主系の障害発生を通知する。アクセス情報変更部１２２ｂは、主系のアクセスしていたディスク＃０のアクセス情報の変更を行う。アクセス情報の変更については、すでに述べた変更と同様である。

【００３５】また、アクセス情報変更部１２２ａ，１２２ｂが、ディスク管理テーブルのアクセス情報を内容に変更がなくとも一定の時間間隔で更新することにより、障害の発生を検出してもよい。アクセス情報変更部１２２ａ，１２２ｂが、アクセス情報を更新する際、更新のタイムスタンプ（日時の記録）も更新される。障害検出部１２４ａ，１２４ｂは、互いに相手のホストシステムのアクセス情報変更部が更新したタイムスタンプをチェックし、タイムスタンプが更新されていないとき、相手

のホストシステムに障害が発生したと判断する。障害発生後の処理は、すでに述べた処理と同様である。

【００３６】以上のように、この実施の形態では、各ホストシステム毎にどのディスクにアクセスするかをアクセス情報としてディスク管理テーブルに設定し、アクセス制御部が、ディスク管理テーブルを参照して、各ホストシステムのディスクへのアクセスを制御するデータストレージシステムについて説明した。アクセス情報変更部は、障害検出部が障害を検出すると、障害を検出したホストシステムのディスクへのアクセスを禁止するようにアクセス情報を変更する。これにより、障害が発生したホストシステムのシャットダウンが、他のホストシステムに影響を与えることを回避する。また、アクセス情報変更部は、障害が検出されたホストシステムがアクセスしていたディスクを他のホストシステムからアクセスできるようにアクセス情報を変更する。これにより、ディスクが利用できない状態になることを回避できる。

【００３７】

【発明の効果】この発明によれば、アクセス制御部が一方の系にＯＳをロードする優先権を与え、他の系からのアクセスをフリーズするので、読み込みの競合が発生せず、正常に読み込みを行うことができる。

【００３８】また、この発明によれば、アクセス制御部が、ディスク管理テーブルを参照して、ホストシステムのディスクへのアクセスを制御する。障害発生時に、アクセス情報を変更し、障害の発生した系からディスクアクセスしないように制御することにより、正常な系に影響を及ぼすことが防止できる。

【図面の簡単な説明】

【図１】　この発明のデータストレージシステムの構成図である。

【図２】　この発明のデータストレージシステムのディスクマネージャ４００の構成図である。

【図３】　この発明のデータストレージシステムの機能ブロック図である。

【図４】　この発明のデータストレージシステムの主系のホストシステムの電源投入からシステム起動完了までの流れ図である。

【図５】　この発明のデータストレージシステムの主系と従系のホストシステムの電源投入からシステム起動完了までのディスクマネージャの動作の流れ図である。

【図６】　この発明のデータストレージシステムの従系のホストシステムの電源投入からシステム起動完了までの流れ図である。

【図７】　この発明のデータストレージシステムのディスクマネージャのＳＣＳＩブロックエリアを示す図である。

【図８】　この発明のデータストレージシステムのシステム起動時のディスクマネージャの動作を示す流れ図である。

【図９】　この発明のデータストレージシステムの構成図である。

【図１０】　この発明のデータストレージシステムのディスク管理テーブルの一例を示す図である。

【図１１】　この発明のデータストレージシステムのアクセス情報変更後のディスク管理テーブルを示す図である。

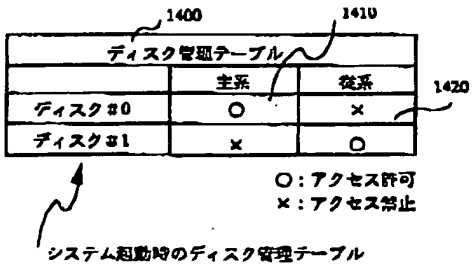
【図１２】　この発明のデータストレージシステムの他のシステム構成図である。

【図１３】　従来の、１つのディスクサブシステムを２つのホストシステムで共有するシステムの構成図である。

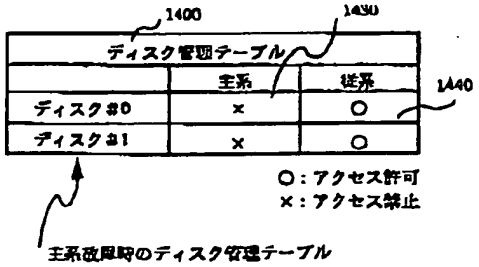
【符号の説明】

１００ａ，１００ｂ　ホストシステム、２００ａ，２００ｂ　バス、３００ディスク、４００　ディスクマネージャ、４１０　ＳＣＳＩバススイッチ、４２０　ＳＣＳＩバスコントロール、４２１　判別部、４２３　アクセス制御部、４２５　読み込み完了情報記憶部、４２６　第１の読み込み完了情報記憶部、４２７　第２の読み込み完了情報記憶部、４３０　メモリ、４３２　ＳＣＳＩブロックエリア、４３８　ディスクマネージャプログラムエリア、４４０　ＭＰＵ、１４００　ディスク管理テーブル。

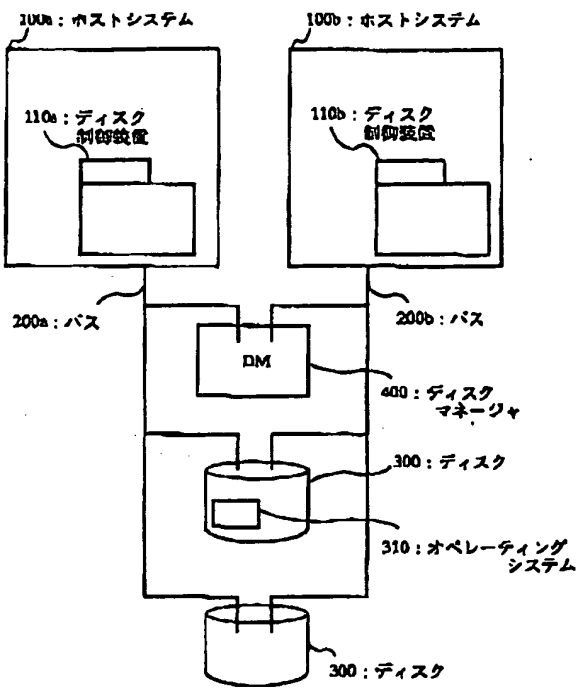
【図１０】



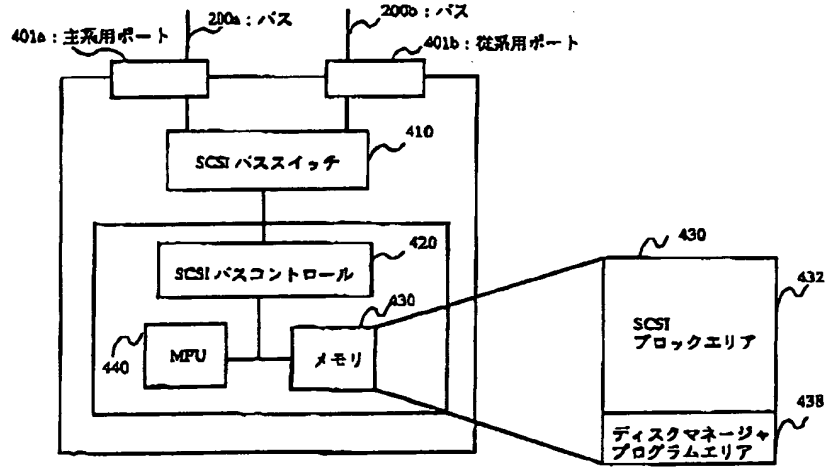
【図１１】



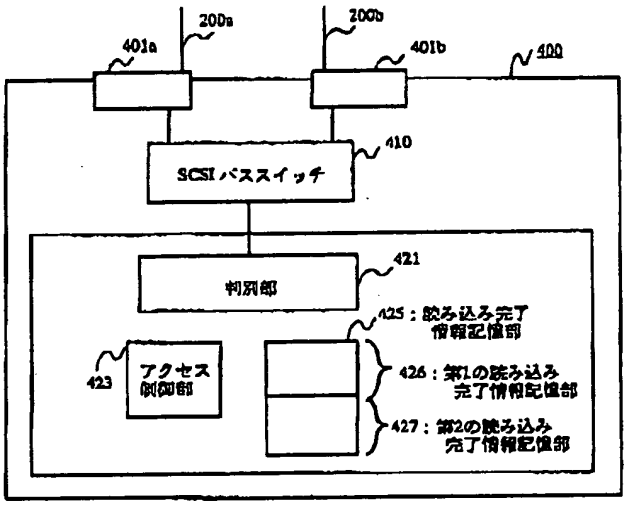
【図１】


100a：ホストシステム
100b：ホストシステム
110a：ディスク制御装置
110b：ディスク制御装置
200a：バス
200b：バス
DM
400：ディスクマネージャ
300：ディスク
310：オペレーティングシステム
300：ディスク

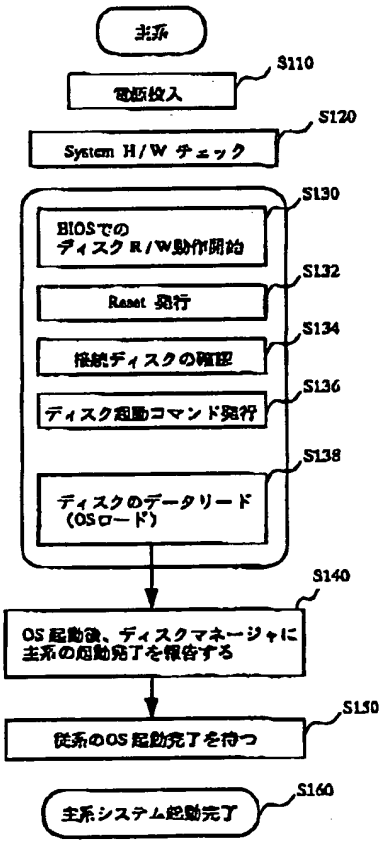
【図２】


401a：主系用ポート
200a：バス
200b：バス
401b：従系用ポート
410
SCSI バススイッチ
420
SCSI バスコントロール
430
MPU
440
メモリ
430
SCSI ブロックエリア
432
ディスクマネージャプログラムエリア
438

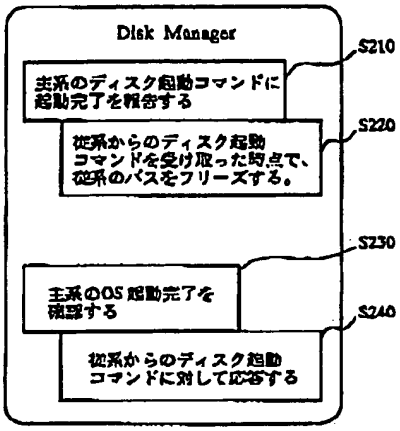
【図３】


401a
200a
200b
401b
400
410
SCSI バススイッチ
421
判別部
425：読み込み完了情報記憶部
423
アクセス制御部
426：第1の読み込み完了情報記憶部
427：第2の読み込み完了情報記憶部

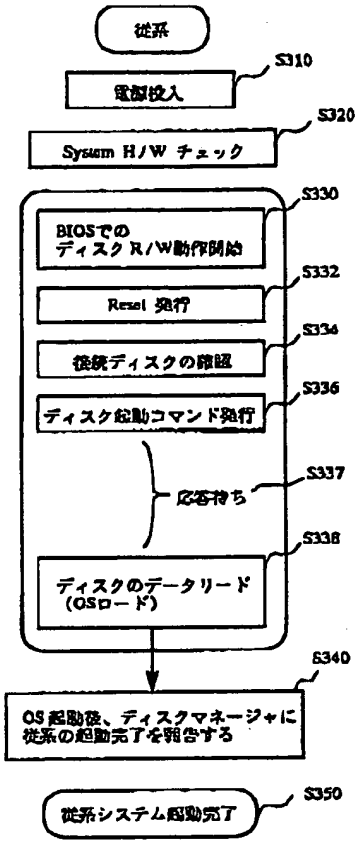
【図４】


主系
S110
電源投入
S120
System H/W チェック
S130
BIOSでのディスク R/W動作開始
S132
Reset 発行
S134
接続ディスクの確認
S136
ディスク起動コマンド発行
S138
ディスクのデータリード（OSロード）
S140
OS 起動後、ディスクマネージャに主系の起動完了を報告する
S150
従系のOS 起動完了を待つ
S160
主系システム起動完了

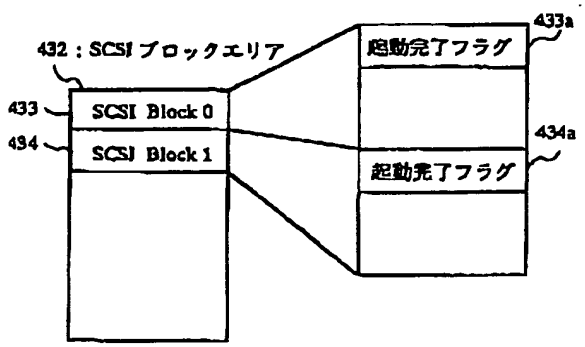
【図５】


Disk Manager
S210
主系のディスク起動コマンドに起動完了を報告する
S220
従系からのディスク起動コマンドを受け取った時点で、従系のパスをフリーズする。
S230
主系のOS起動完了を確認する
S240
従系からのディスク起動コマンドに対して応答する

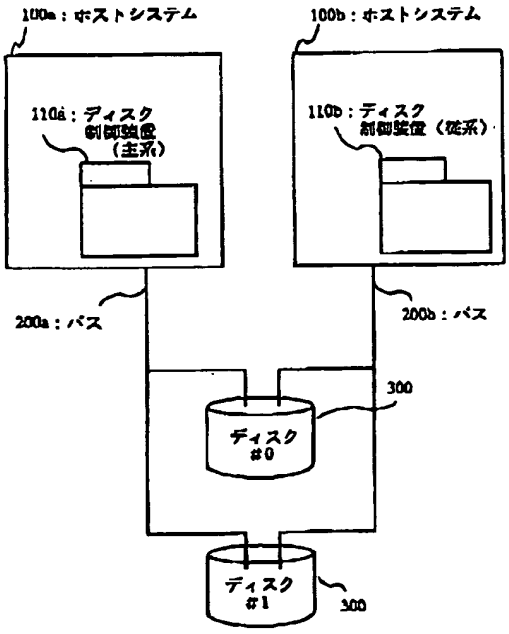
【図６】


従系
S310
電源投入
S320
System H/W チェック
S330
BIOSでのディスクR/W動作開始
S332
Reset 発行
S334
後続ディスクの確認
S336
ディスク起動コマンド発行
S337
応答待ち
S338
ディスクのデータリード（OSロード）
S340
OS起動後、ディスクマネージャに従系の起動完了を報告する
S350
従系システム起動完了

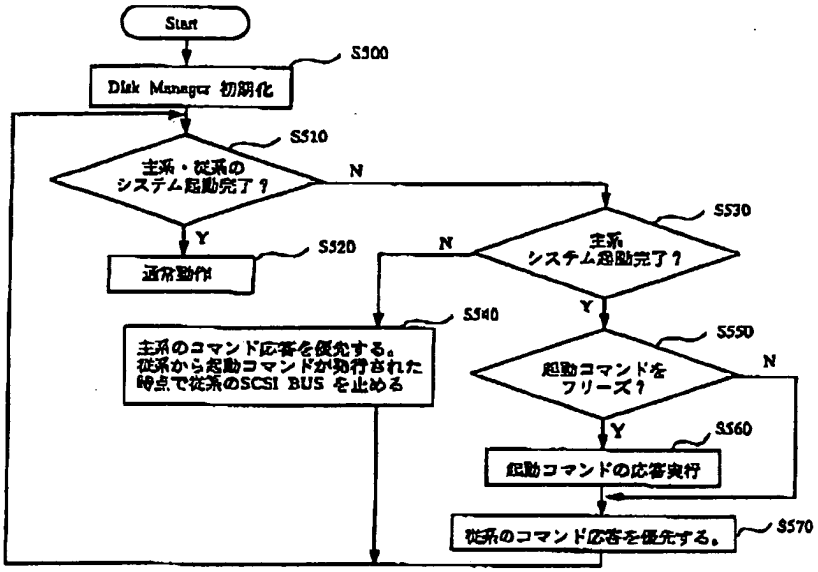
【図７】


432：SCSIブロックエリア
433
434
SCSI Block 0
SCSI Block 1
433a
起動完了フラグ
434a
起動完了フラグ

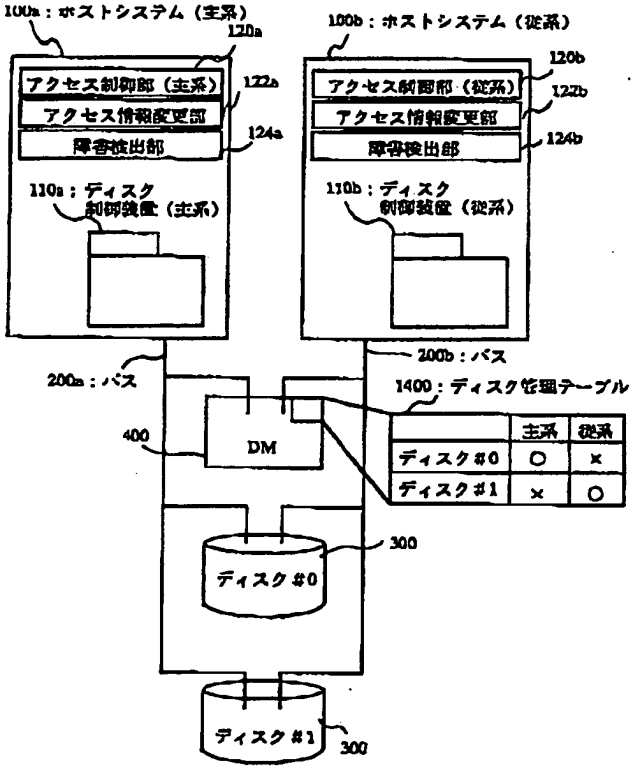
【図１３】


100a：ホストシステム
100b：ホストシステム
110a：ディスク制御装置（主系）
110b：ディスク制御装置（従系）
200a：パス
200b：パス
300
ディスク#0
ディスク#1
300

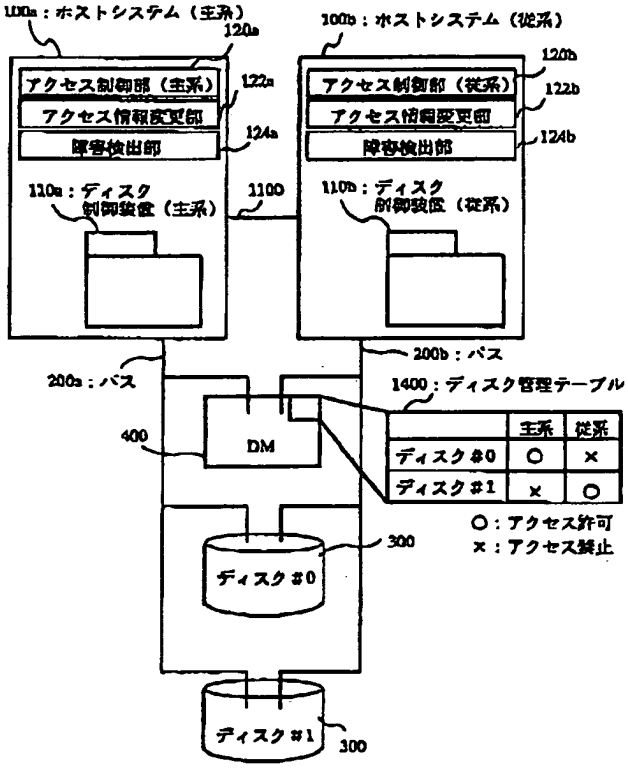
【図８】


Start
S500
Disk Manager 初期化
S510
主系・従系の
システム起動完了？
N
Y
S520
通常動作
S530
主系
システム起動完了？
N
Y
S540
主系のコマンド応答を優先する。
従系から起動コマンドが発行された
時点で従系のSCSI BUS を止める
S550
起動コマンドを
フリーズ？
N
Y
S560
起動コマンドの応答実行
従系のコマンド応答を優先する。
S570

【図９】


100a：ホストシステム（主系）
120a
アクセス制御部（主系）
122a
アクセス情報変更部
124a
障害検出部
110a：ディスク
制御装置（主系）
100b：ホストシステム（従系）
120b
アクセス制御部（従系）
122b
アクセス情報変更部
124b
障害検出部
110b：ディスク
制御装置（従系）
200a：パス
200b：パス
1400：ディスク管理テーブル
400
DM
主系
従系
ディスク#0
○
×
ディスク#1
×
○
300
ディスク#0
ディスク#1
300

【図１２】


100a：ホストシステム（主系）
120a
アクセス制御部（主系）
122a
アクセス情報変更部
124a
障害検出部
110a：ディスク
制御装置（主系）
1100
100b：ホストシステム（従系）
120b
アクセス制御部（従系）
122b
アクセス情報変更部
124b
障害検出部
110b：ディスク
制御装置（従系）
200a：パス
200b：パス
1400：ディスク管理テーブル
400
DM
主系
従系
ディスク#0
○
×
ディスク#1
×
○
○：アクセス許可
×：アクセス禁止
300
ディスク#0
ディスク#1
300